Use of RNA Controls for Control of Micro-array experiments

NIST Meeting 3/28/03

Gianfranco de Feo, Ph.D.

Associate Director, Genomics Collaborations

AFFYMETRIX

RNA Standards in Micro-array expts

- Scope: What do you want from the standards?
  - Quality control
    - Sample
    - Procedure
    - Arrays
  - Normalization
  - Quantitative accuracy
- Goal: Consistency of Data
  - Within platform
  - Across platforms

AFFYMETRIX

Steps to control

- Sample Acquisition and preparation (from sample to RNA)
  - Acquisition
  - Storage and handling
  - RNA isolation
- ‘Target’ Preparation
  - ‘cDNA’ or the like preparation
  - Labeling and amplification
  - Fragmentation and hyb cocktail QC
- Hybridization and Scanning
  - Hybridization, wash and stain
  - Scanning
  - Quantitative metric

AFFYMETRIX

Sample Acquisition and Preparation

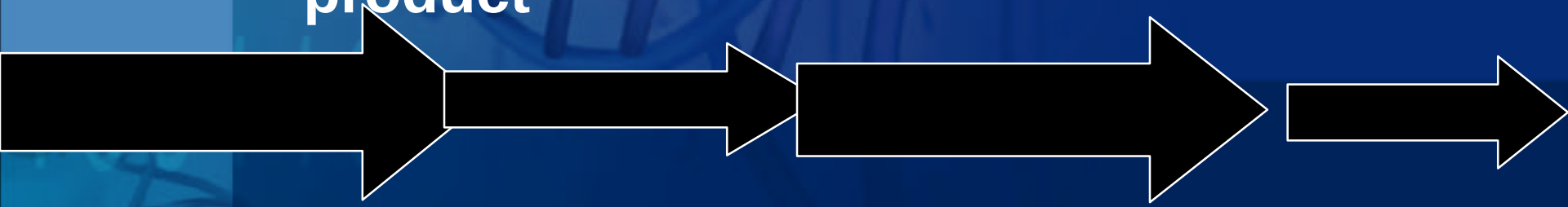
1. Quantity and Quality of RNA
2. Reproducibility: consistency of RNA pool

1. Stabilization of Sample
2. Handling conditions
3. Storage conditions

1. Reproducibility
2. Quantity
3. Quality of RNA

1. Final qualification of sample
2. Quantity and quality of RNA

Target Preparation: From RNA to labeled product



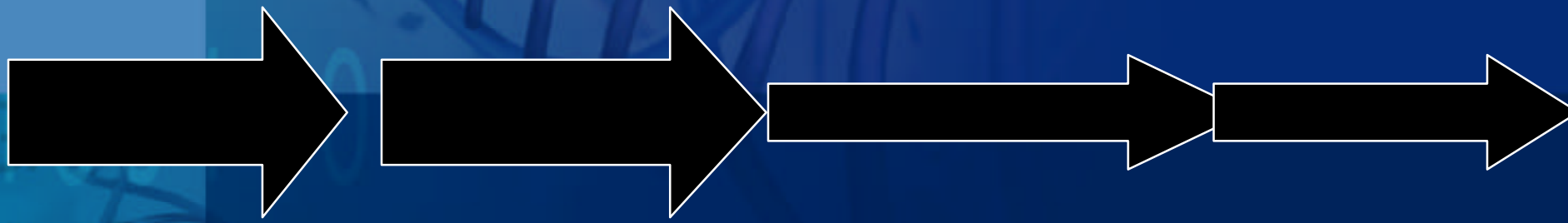
1. Quantity and Quality of cDNA
2. Reproducibility: consistency of cDNA pool

1. Amplification efficiency
2. Representation of final pool
3. Labeling efficiency
4. Consistency

1. Reproducibility
2. Quantity
3. Quality of labeled target

1. Final qualification of sample
2. Quantity and quality of RNA
3. Reproducibility of pool

Array Hybridization and Scanning



1. ‘In process’ QC
2. ‘Probe’ Analysis
3. Consistency of hyb

1. Consistency of signals
2. Standard protocols

1. Reproducibility
2. Image defect tolerances
3. Quality of hybridization

1. Final qualification of results
2. Normalization approach

RNA Controls

- Can be used to assess:
  - Most steps of process (including array)
    - Sample Acquisition?
  - Quantitative Accuracy
  - Normalization strategies
- Types:
  - Poly A+ controls
  - Hybridization controls
    - Pre-labeled transcripts
    - Pre-labeled oligonucleotides
  - Complex sample

AFFYMETRIX

Use of RNA spikes

Type of Spike	‘Synthetic Gene’ RNA Phylogenetically Distant	Synthetic gene RNA Phylogenetically Distant	Pre-labeled RNA
Spike strategy sample	pre-Lysate spikes	Spike into RNA sample	Spike into labeled
	Spike unlabelled Poly A+ containing RNA species		‘Oligo pools Synthetic RNAs’
	Vary spikes according to: sequence, length, concentration		
Monitors	RNA degradation RNA purification	Protocol characteristics	Array assessment Hyb, wash and stain

- Each set of spikes intended to assess a specific part of the procedure
- Not all spikes required, spikes in earlier stages can be used to assess later stages

Poly A+ and pre-labeled Spike Contols

- Advantage: Can control the number, amount and type of spike-ins, in-sample control.
  - Quantitative assessments
  - Dynamic range (linearity of response)
  - Normalization (?)
- Disadvantage
  - Limited complexity
  - Work to generate and characterize
  - QC the spikes
- Scope:
  - quantitatively assess most aspects of pathway:

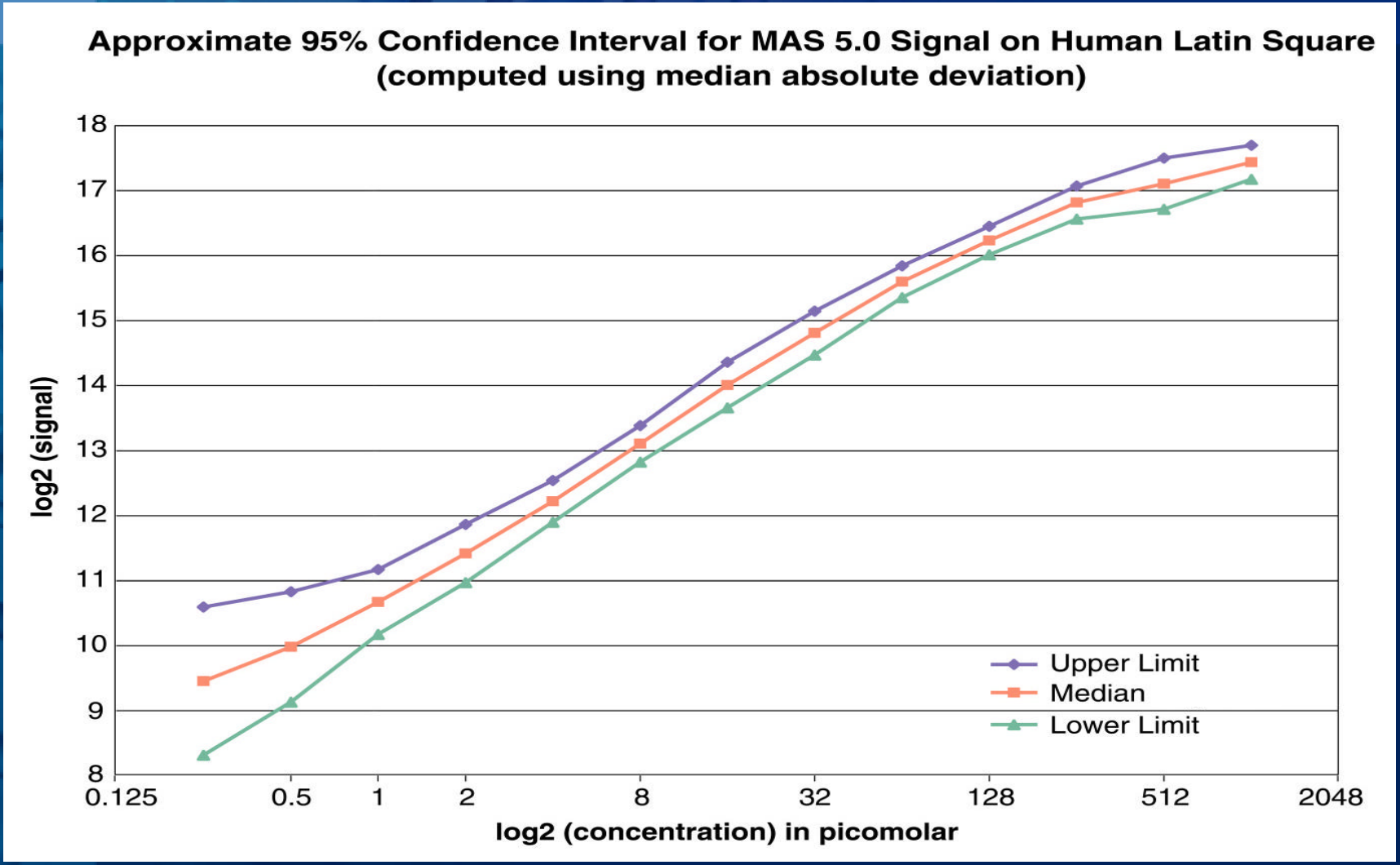
Latin Square Experimental Design

Groups of transcripts
pM concentration

GeneChip experiments

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	0	0.25	0.5	1	2	4	8	16	32	64	128	256	512	1024
2	0.25	0.5	1	2	4	8	16	32	64	128	256	512	1024	0
3	0.5	1	2	4	8	16	32	64	128	256	512	1024	0	0.25
4	1	2	4	8	16	32	64	128	256	512	1024	0	0.25	0.5
5	2	4	8	16	32	64	128	256	512	1024	0	0.25	0.5	1
6	4	8	16	32	64	128	256	512	1024	0	0.25	0.5	1	2
7	8	16	32	64	128	256	512	1024	0	0.25	0.5	1	2	4
8	16	32	64	128	256	512	1024	0	0.25	0.5	1	2	4	8
9	32	64	128	256	512	1024	0	0.25	0.5	1	2	4	8	16
10	64	128	256	512	1024	0	0.25	0.5	1	2	4	8	16	32
11	128	256	512	1024	0	0.25	0.5	1	2	4	8	16	32	64
12	256	512	1024	0	0.25	0.5	1	2	4	8	16	32	64	128
13	512	1024	0	0.25	0.5	1	2	4	8	16	32	64	128	256
14	1024	0	0.25	0.5	1	2	4	8	16	32	64	128	256	512

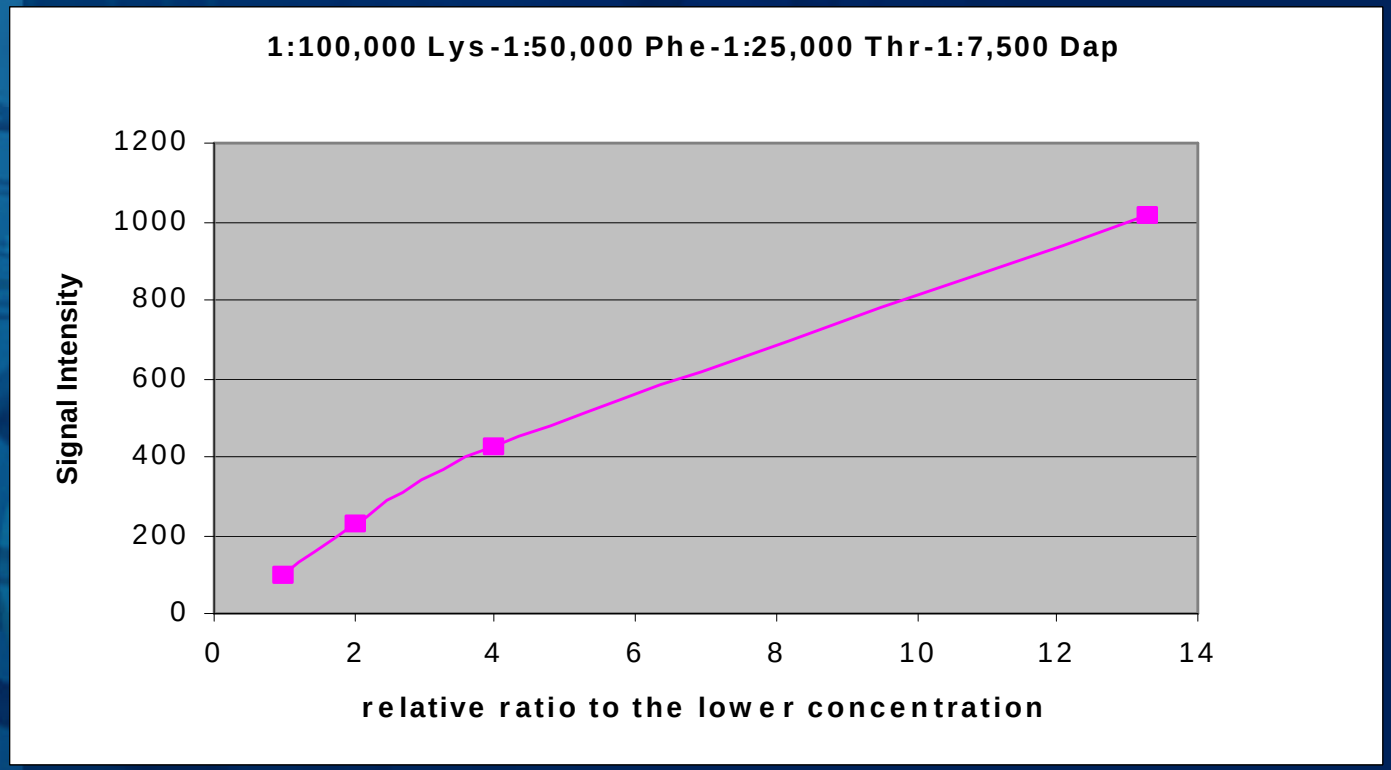
Signal is near-linear and has stabilized variance in the middle range of concentrations



Upper and lower limits refer to 95% confidence intervals

AFFYMETRIX®

Poly A+ Spikes example



Lys	**1:100,000**	**(1)**
Phe	**1: 50,000**	**(2)**
Thr	**1: 25,000**	**(4)**
Dap	**1: 7,500**	**(13)**

What can be assessed:

Sensitivity

Quantitative accuracy and precision

Dynamic range

Complex Sample Discussion

- Advantage: Complexity high and therefore cover larger number of probes on array
  - Process/system QC
  - Cross-platform comparisons
- Disadvantage
  - Thorough characterization of reference sample required. (know the ‘right answer’)
  - Parallel to sample of interest (not ‘in-sample’)
  - QC of the QC tool (lot to lot variability)
  - Pooling different sources of RNA
- Scope:
  - Reference sample for competitive hyb expts.
  - Process QC (used at regular intervals in parallel to samples in expt)

AFFYMETRIX

Complex Sample Example: Coverage

	Mix 1 (5 Tissue Pool)	Mix 2 (10 Cell line Pool)	Single Tissue Source
% Present	51.50%	52.10%	53.50%
False Change	0.07%	0.05%	0.05%

Averages of n=4

Single Tissue source: human Placenta

All expts done with HGU133A arrays

- Mixed source RNA does not seem to increase coverage on the array
- Mixed source RNA harder to reproducibly produce lot to lot
- Could not perform RNA inter-lot reproducibility

Conclusions

- Reasonably complex procedure with many steps to control
  - RNA spikes used to assess and limit variability at most stages
    - Sample acquisition (‘armored’ RNA?)
  - Complex sample for overall process QC
- RNA spikes
  - Use of labeled and unlabeled transcripts
  - Vary according to sequence, length, and concentration
  - Use strategy appropriate for each stage

AFFYMETRIX®